⑫

"END-OF-FISCAL-YEAR" REPORT

Speech Recognition:
Acoustic- Phonetic Knowledge
Acquisition and Representation

Office of Naval Research
Contract N00014-82-K-0727

Covering the Period
1 October 1986 – 30 September 1987

Submitted by:

Victor W. Zue

DTIC
ELECTE
SELECTED
NOV 1 2-1987
E

September 24, 1987

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Research Laboratory of Electronics
Cambridge, Massachusetts 02139

87

"End-of-Fiscal Year" Letter

A. Description of Scientific Research Goals

Our long-term research goal is the development and implementation of speaker-independent continuous speech recognition systems. It is our conviction that the proper utilization of speech-specific knowledge is essential for such advanced systems. Our research is thus directed toward the acquisition of acoustic-phonetic and lexical knowledge, and the application of this knowledge to speech recognition algorithms.

B. Significant Results in the Last Year:

- We continued our investigation into the contextual variations of speech sounds, emphasizing on the role of the syllable in these variations. Our analysis revealed that the acoustic realization of a stop depends greatly on its position within a syllable. In order to represent and utilize this information in speech recognition, we have adopted a hierarchical syllable description that enables us to specify the constraints in terms of an immediate constituent grammar.

- We developed a featured-based framework for phonetic recognition, and implemented a recognition system for semivowels in American English. The recognition process is divided into two stages, detection and classification. Recognition accuracy ranging from 78 to 95% were obtained across different contexts and speakers.

- We continued our efforts to capture the knowledge used by human spectrogram readers and to incorporate it into an expert system. Our emphasis has been on establishing human performance benchmarks, both for auditory perception and spectrogram reading experiments. The results indicate that listeners can correctly identify stops in various environments with accuracy ranging from 85 to 97%. Performance of the spectrogram readers is 10 to 15% lower.

- We refined our system for extracting visual objects from speech spectrograms, using a combination of directional and non-directional edge detectors. Our evaluations of the effectiveness of such representations show that spectrogram readers can recognize speech sounds from such impoverished representations with high accuracy. Also, the recognition system using only the information contained in the objects can achieve comparable performance to that realized using a conventional signal representation. Finally, speech resynthesized from the visual objects is highly intelligible.

- We explored several models of the refractory effect of auditory nerve fibers – that is, the fiber's inability to fire twice in rapid succession. A significant outcome of this study is that the effect contributes a nonlinearity which operates like an automatic gain control. Our tentative conclusion is that an enhancing nonlinearity has evolved in the cochlea so as to nearly counterbalance the compressive refractory effect.
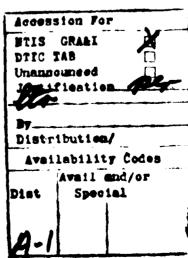
1

- We began work on a spelling recognition task that, taking the 26 letters of the English alphabet as its vocabulary, would recognize continuously spoken letters in the context of spelled words. Our lexical analysis reveals that strong sequential constraints exist for letter strings. Listening and spectrogram reading performances were found to be quite high (98% *vs.* 91%). For those letter pairs that were found to be confusable by humans, we were able to find acoustic parameters that can reliably disambiguate them.

C. Plans for Next Year's Research:

- We will continue to quantify the effect of context on the acoustic realization of phonemes using larger constituent units such as syllables. In addition, we will develop a grammar to describe the relationship between phonemes and acoustic segments, and a parser that will make use of such a grammar for phonetic recognition and lexical access.

- We will begin a study to quantify the effect of various factors, including speaking rate, on the temporal structure of speech, and to develop a durational model that takes these factors into account.

- We will explore alternative phonetic recognition and lexical access strategies using neural nets, and also compare neural net approaches to stochastic modelling approaches such as hidden Markov modelling.

- We will extend and refine our experimental phonetic recognition system based on an expert system that mimics a spectrogram reader. Declarative knowledge for stops in singletons and selected clusters will be embodied in the system.

- We will begin to explore a speech recognition strategy based on distinctive feature theory. We will place our emphasis on identifying and implementing acoustic property detectors for some of the important distinctive features, and on determining lexical access strategies based on the distinctive features.

- We will explore alternative language modelling strategies that embed traditional linguistic descriptions into a stochastic framework, so that more habitable grammars can be developed for speech understanding.

D. List of Presentations: Please see enclosure 2.

E. List of Technical Reports: NONE.

F. List of Publications: Please see enclosure 2.

G. List of Honors/Awards: NONE.

2

H. Participants:

Principal Investigator
Victor W. Zue

Research Staff
Dennis Klatt
David Kaufman
Niels Lauritzen
Michael Phillips
Stephanie Seneff

Graduate Students
Nancy Daly (S.M. degree granted June, 1987)
Susan Dubois
Carol Espy-Wilson (Ph.D. degree granted June, 1987)
James Glass
Rob Kassel
Lori Lamel
Hong Leung
John Pitrelli
Mark Randolph
Timothy Wilson (S.M. degree granted June, 1987)

Undergraduate Students
Charles Jankowski
Amy Lim
Hirak Mitra
Andrew Shaw
Sean Tierney (S.B. degree granted June 1987)
David Whitney

I. Other Sponsored Research:

- Title: Acoustic-Phonetics Based Speech Recognition
  Sponsor: Naval Electronic Systems Command (DARPA)
  Amount: $2,018,533.00
  Contract Period: 8 February 1985 - 31 January 1989

- Title: Speech Database Development
  Sponsor: Naval Electronic Systems Command (DARPA)
  Amount: $501,194.00
  Contract Period: 20 May 1985 - 19 August 1987

- Title: Tools for Speech Analysis and Research
  Sponsor: Space and Naval Warfare Systems Command (DARPA)
  Amount: $641,526.00
  Contract Period: 27 June 1985 - 26 June 1987

J. Computer Net Address: zue%speech@mc.lcs.mit.edu

List of Publications/Reports/Presentations

1. Papers Published in Refereed Journals and Conference Proceedings:

- Randolph, M. A., and V. W. Zue, "The Influence of Phonetic Context on the Acoustic Properties of Stops," 112th Meeting of the Acoustical Society of America, Anaheim, CA, Dec. 1986.

- Randolph, M. A., and V. W. Zue, "The Role of Syllable Structure in the Acoustic Realizations of Stops," *Proc. 11th International Congress of Phonetic Sciences*, 1987, 36.2.1-36.2.4.

- Espy-Wilson, C. Y., "A Semivowel Recognition System," *Proc. 11th International Congress of Phonetic Sciences*, 1987, 95.4.1-95.4.4.

- Leung, H. C., and V. W. Zue, "Two-Dimensional Characterization of the Speech Signal and its Potential Applications to Speech Processing," to be presented at *First International Conference on Communication Technology*, 1987.

Theses:
2. ~~XXXXXXXXXX XXXXXXXXX~~

- Espy-Wilson, C. Y., "An Acoustic-Phonetic Approach to Speech Recognition: Application to the Semivowels," Ph.D. thesis, Massachusetts Institute of Technology, May, 1987.

- Daly, N. A., "Recognition of Words from their Spellings: Integration of Multiple Knowledge Sources," S.M. thesis, Massachusetts Institute of Technology, May, 1987.

3. Presentations/

a. Invited

- Zue, V. W., "How to Incorporate Knowledge into Automatic Speech Recognition," NYNEX Science and Technology Symposium on Speech Processing, Boston, MA, and New York, NY, Mar. 16-17, 1987.

- Zue, V. W., "Automatic Speech Recognition: Trends and Applications," US West Advanced Technologies Future Technologies Forum, Denver, CO, June 29-30, 1987.

b. Contributed

4. Books (and sections thereof)

LIST OF AWARDS

Name of Person Receiving Award	Recipient's Institution	Name of Award	Sponsor of Award
NONE			

Enclosure (3)

PUBLICATIONS/PATENTS/PRESENTATIONS/HONORS REPORT
(NUMBER ONLY)

Papers Submitted to Refereed Journals (and not yet published): 1

Papers Published in Conference Proceedings: 3

Books (and sections thereof) Submitted for Publication: 0

Books (and sections thereof) Published: 0

Patents Filed: 0

Patents Granted: 0

Invited Presentations at Topical or Scientific/Technical Society Conferences: 3

Contributed Presentations at Topical or Scientific/Technical Society Conferences: 0

Honors/Awards/Prizes: 0

Theses: 2

Number of Graduate Students: 9

Number of Post Docs: 1

Enclosure (4)

DISTRIBUTION LIST

	DODAAD Code	
Head Information Sciences Division Office of Naval Research 800 North Quincy Street Arlington, Virginia 22217	N00014	(1)
Administrative Contracting Officer E19-628 Massachusetts Institute of Technology Cambridge, Massachusetts 02139		(1)
Director Naval Research Laboratory Washington, D.C. 20375 Attn: Code 2627	N00173	(1)
Defense Technical Information Center Bldg. 5, Cameron Station Alexandria, Virginia 22314	S47031	(12)

END

FEb.

1988

DTIC